Diffusion-Based Reinforcement Learning for Edge-Enabled AI-Generated Content Services

Year: 2023 | Citations: 123 | Authors: Hongyang Du, Zonghang Li, Dusist Niyato

Abstract

As Metaverse emerges as the next-generation Internet paradigm, the ability to efficiently generate content is paramount. AI-Generated Content (AIGC) emerges as a key solution, yet the resource-intensive nature of large Generative AI (GAI) models presents challenges. To address this issue, we introduce an AIGC-as-a-Service (AaaS) architecture, which deploys AIGC models in wireless edge networks to ensure broad AIGC services accessibility for Metaverse users. Nonetheless, an important aspect of providing personalized user experiences requires carefully selecting AIGC Service Providers (ASPs) capable of effectively executing user tasks, which is complicated by environmental uncertainty and variability. Addressing this gap in current research, we introduce the AI-Generated Optimal Decision (AGOD) algorithm, a diffusion model-based approach for generating the optimal ASP selection decisions. Integrating AGOD with Deep Reinforcement Learning (DRL), we develop the Deep Diffusion Soft Actor-Critic (D2SAC) algorithm, enhancing the efficiency and effectiveness of ASP selection. Our comprehensive experiments demonstrate that D2SAC outperforms seven leading DRL algorithms. Furthermore, the proposed AGOD algorithm has the potential for extension to various optimization problems in wireless networks, positioning it as a promising approach for future research on AIGC-driven services.